Technical Disclosure Commons

Defensive Publications Series

June 2020

MINIMAL DELAY WIFI DISPLAY SHARING WITH USER INPUT BACK CHANNEL IN VIDEO CONFERENCING

Harald Dankworth

MINIMAL DELAY WIFI DISPLAY SHARING WITH USER INPUT BACK CHANNEL IN VIDEO CONFERENCING

AUTHORS:
Harald Dankworth

ABSTRACT

Techniques are provided herein for a WiFi Display bridge implemented in telepresence endpoints. This may enhance collaboration technology by upgrading collaboration presentation devices to a remote desktop without requiring software installation. The far-end collaboration presentation devices receive high-fidelity and low-latency audio and video screen captures of the presenting device and have the capability to control the mouse and keyboard.

DETAILED DESCRIPTION

WiFi Display is a technical specification by the WiFi Alliance, which also provides the Miracast certification. WiFi Display enables wireless sharing to a telepresence conference without installing any software or requiring any High-Definition Multimedia Interface (HDMI) cables. Some telepresence vendors have support for WiFi Display in their videoconferencing endpoints. This enables wireless sharing to the local endpoint which can then forward the screen share to remote participants. Typically, the wireless share would be decoded on the local endpoint and then encoded again for transmission on the video conferencing presentation channel. This can contribute to delay beyond the maximum permissible delay. For example, the delay must be within 250 ms to pass Miracast certification.

Today, the capability does not exist to use a WiFi Display source as a remote desktop for a presentation with user interaction in a telepresence scenario. In particular, when screen sharing wirelessly from a WiFi Display device to a videoconferencing endpoint, the remote conferencing participants are able to see the screen share, but not interact with it. The collaboration would be more efficient if the remote participants were able to use the screen sharing presentation as a remote desktop. In order to achieve this, the remote participants must be able to provide user inputs on their devices, and the user inputs must be forwarded to the presentation source. Moreover, a good user experience requires minimal delay. Accordingly, techniques are described herein to use the WiFi Display presentation source as a remote desktop in a telepresence system without requiring any software installation on the presentation source. This is achieved using reduced delay and a user input backchannel over telepresence.

With respect to reduced delay, the local telepresence system receives the capability set of the remote participants by existing means (e.g., Session Initiation Protocol (SIP), H.323, proprietary signaling, etc.). This capability set is used in the Real Time Streaming Protocol (RTSP) capability exchange with the WiFi Display presenting device, instructing the presenting device to encode within the capabilities of the remote participants. Real-time Transport Protocol (RTP) packet loss indication or a signaled fast update to the local telepresence system is forwarded to the WiFi Display over RTSP. The Real-time Transport Control Protocol (RTCP) receiver reports on the local presence system are aggregated and forwarded to the WiFi Display over RTCP, enabling the WiFi Display to adjust its video rate control to the channel conditions. If, for some reason, it is not possible to use the WiFi Display encoding directly, the local telepresence system may decode and re-encode before sending network communications to the remote telepresence participants, albeit at the cost of longer latency.

With respect to the user input backchannel over telepresence, remote telepresence systems can have a user input (e.g., a touch screen where the presentation is rendered). The captured user input events are sent from the remote device to the local telepresence system with application-defined RTCP packets. The local telepresence system forwards the user input received via RTCP to the WiFi Display using the standardized User Input Back Channel (UIBC). If the remote telepresence system does not have an embedded touch screen, the endpoint could act as a WiFi Display source and forward the presentation to a local WiFi Display capable touch screen or directly to a Windows® laptop acting as a WiFi Display sink. The telepresence systems would then effectively act as a bridge between the local presentation WiFi Display source and the remote WiFi Display sink touch display.

Unlike current WiFi Display sink implementations, the UIBC is extended/bridged to the far-end participants in video conferences. This is outside the scope of the current WiFi Display technical specification, in which the UIBC is terminated in the host telepresence system (WiFi Display sink). An important quality metric for remote desktop implementations is latency, and for the purposes of WiFi Display, the main contributors to latency is video encoding and decoding. By propagating capabilities and fast update requests from the far-end participants in a video conference to the WiFi Display source, one extra cycle of decoding and encoding is avoided in the host telepresence system.

WiFi Display has various rate-adaption schemes. In accordance with techniques described herein, the WiFi Display source may be signaled to in such a manner that the encoding can be used directly as a "duo channel" video encoding in the video conference directly. Thus, transcoding may be avoided by controlling the maximum frame-rate, frame-size, and bitrate. Packet loss in the duo channel is handled by instructing the WiFi Display source to encode an intra frame.

Figure 1 below illustrates an example system. As shown, a Windows laptop is presenting over a WiFi access point to the local telepresence system, which has support for a WiFi Display sink. The Android® device is presenting over WiFi Direct to the local telepresence system. The local telepresence system has a touch screen and can use the standard WiFi Display UIBC. The local telepresence system has an embedded touch display and is in a call with a remote telepresence system that is presenting to an external touch display with WiFi Display support. The WiFi Display RTP stream is forwarded unmodified if the presenting device is generating a stream compliant with the capabilities of the remote systems. Otherwise, it must be transcoded first. In the case where the remote participant uses an external display, the endpoint must act as a WiFi Display source. This effectively creates a WiFi Display tunnel that enables upgrading the duo presentation channel to a "duo remote desktop channel."

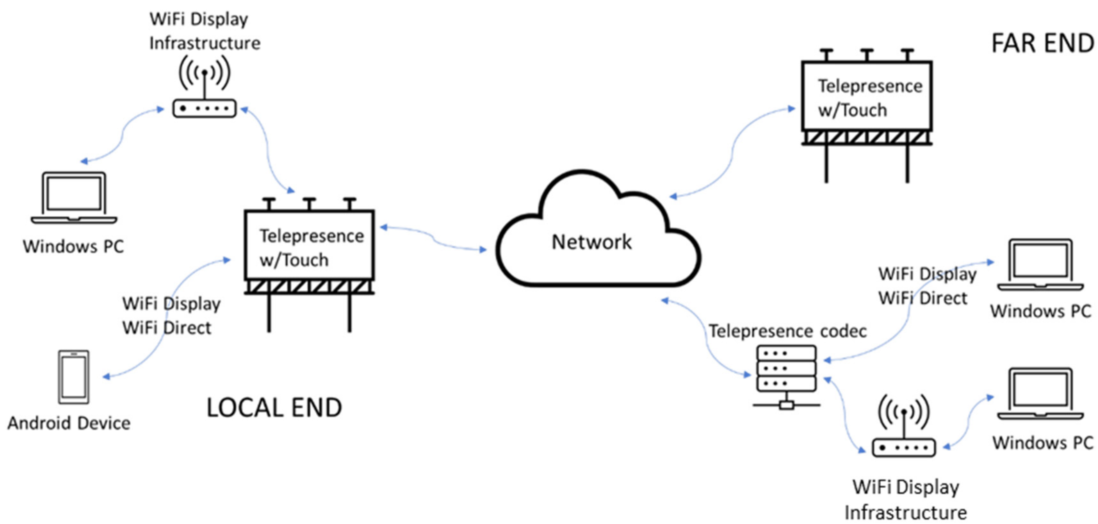

Figure 1

Figure 2 illustrates how to bridge the far-end with the WiFi Display source. Far-end capabilities such as audio codec, audio channels, audio sampling frequency, video codec, maximum frame size, maximum framerate, and maximum bitrate are signaled to the WiFi Display source in order to avoid an extra decoding and encoding cycle in the local telepresence system. RTCP Packet Loss Indication (PLI) from the far-end is translated into an RTSP fast update request for the WiFi Display source. RTCP application specific messages are used to transport user input events from the far-end to the local-end. These are then translated into WiFi Display UIBC events which are received at the WiFi Display source.

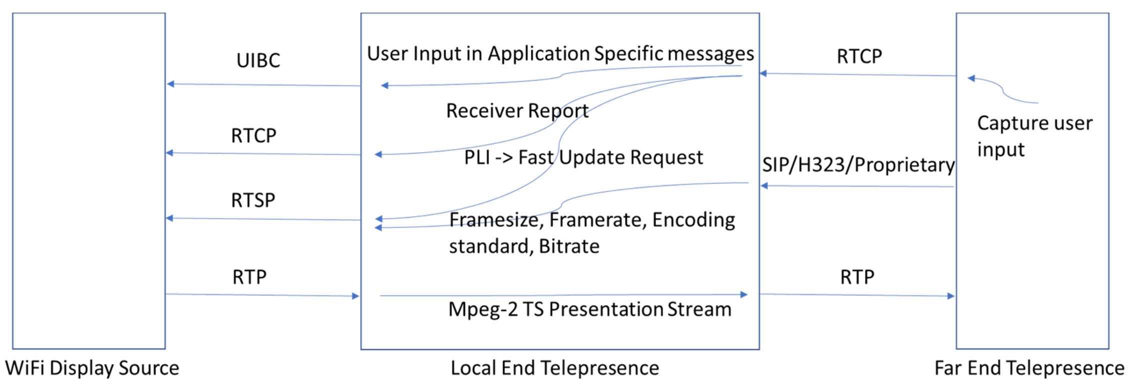

Figure 2

In summary, techniques are provided herein for a WiFi Display bridge implemented in telepresence endpoints. This may enhance collaboration technology by upgrading collaboration presentation devices to a remote desktop without requiring software installation. The far-end collaboration presentation devices receive high-fidelity and low-latency audio and video screen captures of the presenting device and have the capability to control the mouse and keyboard.